Accelerating Drug Development Through the Use of Combination Phase II/III Designs

Jonathan R. Smith, Ph.D.

June 19th, 2002, DIA Annual Meeting, Chicago

Outline of Presentation

- Adaptive design introduction
- Potential uses of Phase II/III combination designs
- Example of **2** look Phase II/III design
  - Type I error control
  - Power assessment
  - Timing of Look #1
- Example of **3** look Phase II/III design
- Comparison vs. separate Phase II & Phase III
- Future Work
- Conclusions

Some Types of Adaptive Designs

	Incorporated here
1. Stop for futility	Yes
2. Stop early for efficacy	Yes
3. Sample size re-estimation	Future
4. Add in new arms	No
5. Change randomization ratio	No
6. Change primary endpoint(s)	No
7. Change test statistic	No

This talk describes combination Phase II/III designs incorporating 1-2

Situations Where Adaptive Designs are Most Useful

The following also applies to interim analysis in general:

1. Follow-up short relative to duration of enrollment
2. Short time from LPV to Interim Analysis decision
3. Randomization via IVR if arms can be dropped

Note: If follow-up is long but onset of action is rapid can often overcome #1 by using result on primary endpoint (or surrogate) but from early visit

Phase II/III Combination Trial - Situations Where this could be Useful

- Range of doses will likely cover optimal dose
  - so that arms will not need to be added
- At most 5 doses still under consideration
- Major safety concerns not likely to apply to very many doses
  - otherwise separate Phase II is probably preferable

Phase II/III Combination Trial - Situations Where this could be Useful, cont'd

- Certain Single Study Submissions
  - Fast Track (if Phase II, III endpoints are the same)
  - New stage of disease, or closely related disease
  - New patient population
  - New combination therapy
  - Orphan indication or other rare disease

Phase II/III Combination Trial - Situations Where this could be Useful, cont'd

- Where sponsor would otherwise carry out Phase II, Phase III #1, Phase III #2 in sequence due to limited funds
- Methodology may also sometimes be useful in place of a multi-armed Phase III trial
  - e.g., in place of past Phase III trial that included 5 doses (50-fold range) and Placebo

Aims of Phase II/III Combination Studies Considered Here:

- Combine dose selection and confirmatory stages
  - start with 2-5 doses + placebo
- Not stop for success in first 50% of study
  - due to safety database needs
- Stop for success as early as possible once we have enough patients for safety
- Stop study early if all doses are ineffective

Phase II/III Combination Design - Two Look Case

1. At start randomize to PL, $D_1,\ldots D_k$
2. At Interim
   - Choose "best" dose D_{BD}
   - *Decision rule specified in protocol & administered by independent group*
   - Stop if all D_i futile
   - Randomize to PL, D_{BD} from now onwards
3. At Final
   - Test D_{BD} vs. PL at level α_2
   - Efficacy demonstrated if test statistic $\geq Z_{1-\alpha 2}$

Two Look Case, cont'd

- Suppose data is normally distributed
- Z_{iDj} = test statistic at look #i for D_j vs. PL
- Test at 1-sided α_i for i=2 (onwards)
- $< Z_{1-\alpha 0}$ is futility decision rule used to stop study at look #1 (with corresp. CP)
- n_1 per group at look #1, n_2 extra per group at look #2
  - also generalized to allow unequal #s per group

Two Look Case- Type I Error

Type I error is given by

$$\sum_{j=1}^{k} P(Z_{1Dj} \geq Z_{1-\alpha_0} \cap Z_{2Dj} \geq Z_{1-\alpha_2} \cap D_j \text{ selected })$$

Note: this applies whatever decision rule is used to select D_j

Two Look Case - Type I Error, cont'd

Suppose decision rule at look #1 is to choose D_j corresponding to highest Z_{1Dj}

Type 1 error can be shown to be given by

$$\int_{z_{1-\alpha_0}}^{\infty} \Phi\left(\sqrt{\frac{n_1}{n_2}}v - Z_{1-\alpha_2}\sqrt{\frac{n_1+n_2}{n_2}}\right)$$

$$* \, P\{\max(Z_{1D1}...Z_{1Dk} = v)\} \, dv$$

See also Simon et al (1994), Hsu et al (1997), Todd & Stallard (2001)

Calculation of Critical Alpha levels

- Equate previous equation to $\alpha = 0.025$ (1-sided)
- 2d Integral evaluated numerically making use of results from Dunnett (1955)
- For given n_1, n_2, k, α, solve for α_2
- Could increase α_2 even further (as Tsong et al, 1997) by allow for unspent T1E resulting from stopping study due to futility
  - For now, not made use of this in case sponsor decides to override DSMB recommendation to stop and continues with the two-arm trial for stage 2

Critical Alpha levels

k	10%	20%	30%	50%	99.9%	Dunnett's alpha
1	0.02500	0.02500	0.02500	0.02500	0.02500	0.02500
2	0.01919	0.01751	0.01645	0.01510	0.01350	0.01348
3	0.01667	0.01443	0.01306	0.01136	0.00944	0.00941
4	0.01517	0.01266	0.01115	0.00933	0.00733	0.00731
5	0.01414	0.01147	0.00990	0.00803	0.00603	0.00601

Determination of Power

- This is analytically more complex and so for now is determined by simulation (100,000 runs for each example)
- Example with k=4 active arms
  - $\mu/\sigma = (0, 0.07, 0.14, 0.21, 0.22)$
  - $\mu/\sigma = (0, 0.06, 0.12, 0.18, 0.24)$
  - $\mu/\sigma = (0, 0.03, 0.11, 0.20, 0.23)$
  - $\mu/\sigma = (0, 0.08, 0.16, 0.24, 0.16)$
- # patients in trial fixed at $1500 = 5n_1 + 2n_2$

Power and Timing of Interim #1

n_1	n_2	Interim #1	α_2	Power			
				Case 1	Case 2	Case 3	Case 4
30	675	4.3%	0.01788	85.8%	82.4%	82.9%	84.8%
60	600	9.1%	0.01550	88.4%	86.0%	87.4%	**87.2%**
90	525	14.6%	0.01381	**89.4%**	**87.8%**	**89.4%**	**88.1%**
120	450	21.1%	0.01247	**89.8%**	**88.7%**	**89.9%**	**88.3%**
150	375	28.6%	0.01133	**89.3%**	**88.6%**	**89.6%**	**87.9%**
240	150	61.5%	0.00865	83.3%	83.9%	84.2%	83.7%

Timing of Interim #1 with k=4

- When determining timing of Interim #1 seek to balance
  - (a) Need for high power in current study
  - (b) High chance that "best dose" is selected at interim #1
  - (c) Time of interim is late enough for good dose-response information to be obtained
- In examples with k=4, having interim at 20% approximately, gave highest power.
  - 20%-30% may be preferable when allow for (b)-(c)

Extension to 3 or more Looks

1. At start randomize to PL, $D_1,\ldots D_k$
2. At Interim #1
   - Choose "best" dose D_{BD}
   - Stop if all D_i futile
   - Randomize to PL, D_{BD} from now onwards
3. At Interim #i (i>1) and Final (look #r)
   - Test D_{BD} vs PL at level α_i
   - Stop trial for efficacy if test statistic $\geq Z_{1-\alpha i}$

Three Look Case - Type I Error

Type I error in normal case:

$$\sum_{j=1}^{k} P(Z_{1Dj} \geq Z_{1-\alpha_0} \cap Z_{2Dj} \geq Z_{1-\alpha_2} \cap D_j \text{ selected})$$

$$+ \sum_{j=1}^{k} P(Z_{1Dj} \geq Z_{1-\alpha_0} \cap Z_{2Dj} < Z_{1-\alpha_2} \cap Z_{3Dj} \geq Z_{1-\alpha_3} \cap D_j \text{ selected})$$

Note: this expression applies whatever decision rule is used to select D_j

Three Look Case - Type I Error, cont'd

- Suppose, as before, decision rule at look #1 is to choose D_j with highest Z_{1Dj}
- Type 1 error can be expressed as a multivariate normal probability, evaluated numerically
- Dimensionality can be reduced by allowing for independent increments (Todd & Stallard, 2001)
- For given $n_1,\ldots n_r$, k, α, and spending function (relating α_2, α_3, ... α_r) solve for α_r

Three Look Case - Example

Suppose that:

- Pocock-like α-spending function is used, i.e., $Z_{1-\alpha 2} = Z_{1-\alpha 3}$
- No look for early efficacy at Interim #1
  - may want to incorporate extreme Haybittle-Peto like bound at interim #1, using $Z_{1-\alpha 1} = 6.0$
- Looks for efficacy at 75%, 100%
- Example values of μ/σ as before, with k=4
- # patients in trial fixed at $1500=5n_1+2n_2+2n_3$

Power and Timing of Interim #1 - 3 Look Case

Interim #1	$\alpha_2 = \alpha_3$	Overall Power			
		Case 1	Case 2	Case 3	Case 4
4.0%	0.01159	84.5%	80.9%	81.6%	83.5%
9.1%	0.00974	87.3%	85.0%	86.3%	85.8%
13.5%	0.00878	**88.2%**	86.4%	**87.9%**	86.4%
20.1%	0.00777	**88.4%**	**87.5%**	**88.4%**	**87.1%**
25.9%	0.00714	**87.8%**	**87.5%**	**88.5%**	**86.7%**
32.5%	0.00659	86.9%	**87.1%**	**87.7%**	86.4%
52.4%	0.00552	83.2%	83.9%	83.9%	83.3%

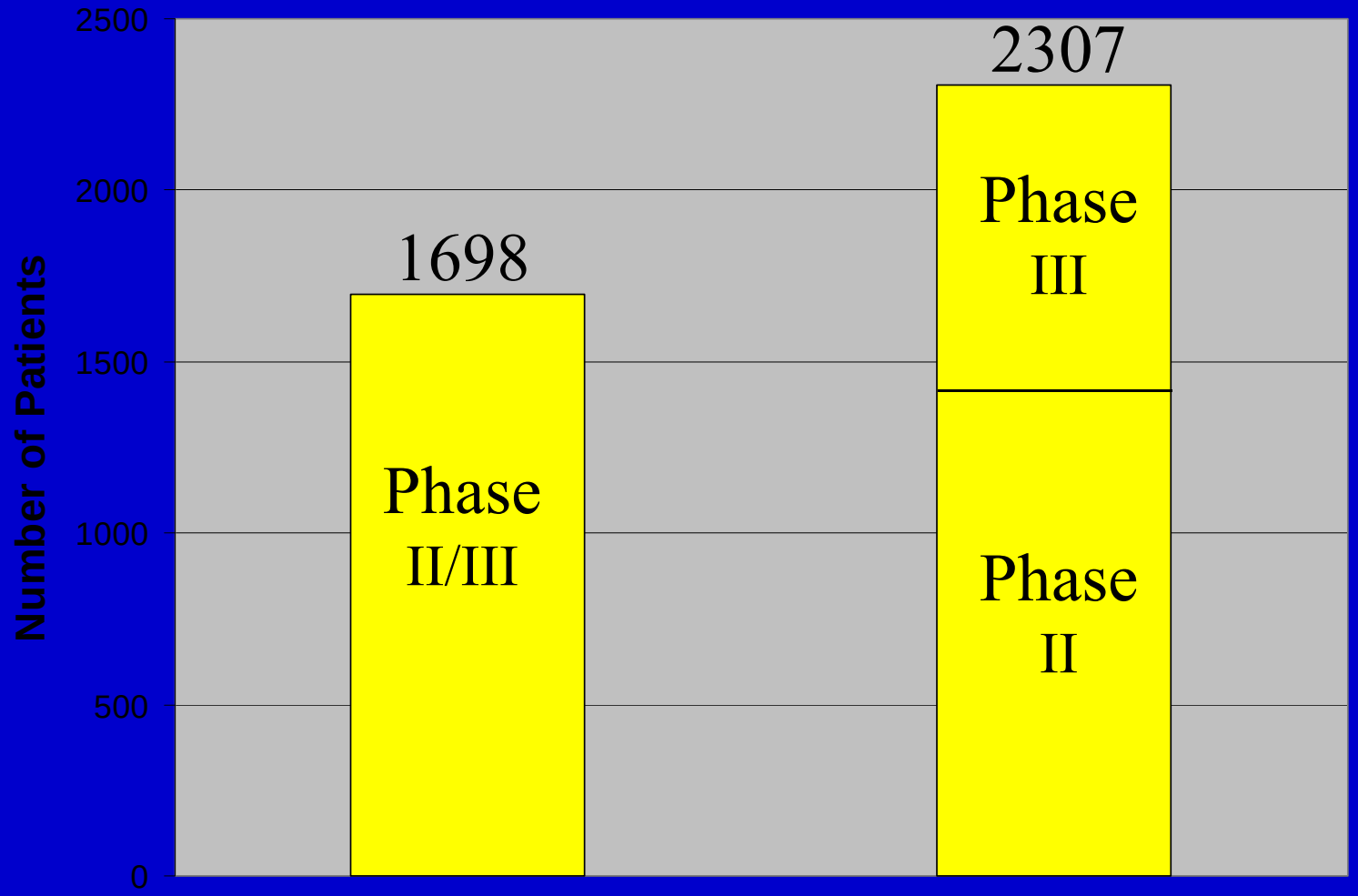
Sample Size Calculations for k=4 - Separate Phase II & Phase III Trials

- Assumptions
  - Interim #1 at 20% in Phase II/III, i.e., $n_2 = 4\ n_1$
  - Equal replication
  - $\mu/\sigma = (0, 0.07, 0.14, 0.21, 0.22)$
- For separate trials we require 2307 patients
  - 287 per arm (1435 total) for 5 arm Phase II
    based on Williams' test with 80% power
  - 436 per arm (872 total) for 2 arm Phase III
    based on $\Delta/\sigma = 0.22$, 90% power

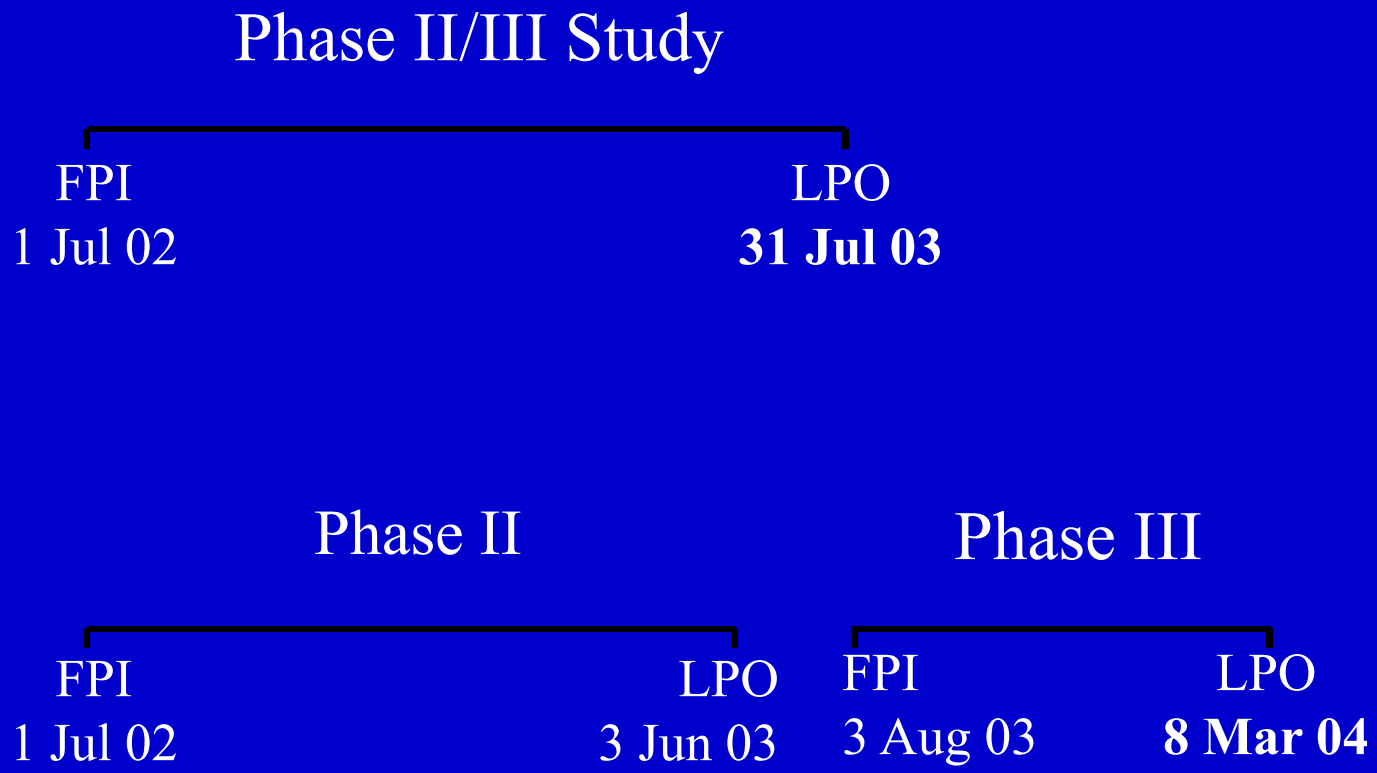
Sample Size Calculations for k=4 - Combined Phase II/III Trial

- Phase II/III (90% power) requires 1698 patients
  - n_1=118 in each of 5 groups prior to Interim #1
  - n_2=472 extra in PL and selected dose group
  - $1534 = 2\ n_2 + 5\ n_1$
  - Also have 164 patients enrolled between last interim patient and implementation of dropping of their arm
    - Assumes 12m enrollment (33 per week), primary endpoint at 30 days, 4w from LPV to drop 3 arms

Relative Number of Patients



Comparison of Timelines



Phase II/III Combination Trial vs. Separate Phase II & Phase III

Benefits of Phase II/III trial in this example:

- **609 Fewer Patients** (2307 - 1698) needed as use Phase II data in final analysis
- **7.3 Months Saved**
  - 18.5 weeks due to enroll 609 fewer patients (33/w)
  - >3 months between LPI study #1 and FPI study #2

Assuming 30d duration in Phase II + 4w from LPO to dose selection + 4w to start up 2nd study & gear up enrollment again

Further Work

- More extensive evaluation of current Phase II/III design approach
  - k = 2, 3, 5 & broader sets of Δ/σ, etc.
  - further comparison vs separate Phase II & Phase III
- Modified decision rules
  - Allow for shape of dose-response curve in decision rule
  - Extend calculations to take account of chance that D_j has safety problems, where dose with maximal Z_{1Dj} may not then be chosen
    - T1E still controlled
    - Impact on power needs assessment

Further Work, cont'd

- Allow two "best" doses (& PL) to be kept after interim #1
- Incorporate sample size re-estimation
  - unblinded, extending Liu & Chi (2001), Cui et al (1999), or Bauer & Kohne (1994)
  - blinded, extending Gould & Shih (1991, 1998)

Conclusions

- Methodology presented allows combination of dose-finding and confirmatory stage within one study
- Type 1 error is controlled exactly at 0.025
- Allows early stopping if all doses are clearly ineffective
- Can be combined with any alpha-spending function to enable stopping as early as possible, subject to meeting safety database needs

Conclusions, cont'd

- Allows for dose selection based on Phase III primary endpoint
- Preliminary results show that if we start with as many as 4 doses of test drug, then having dose selection at 20% - 30% gives
  - high power
  - low chance of continuing with a sub-optimal dose
  - adequate dose-response information

Conclusions, cont'd

- Preliminary results indicate that, where this approach is consistent with needs of program, it can:
  - Cut Drug Development Time
    7.3 Months Time Saving in Example
  - Cut Costs by Reducing Number of Patients
    609 Fewer Patients Needed in Example